

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Katz, Leonard
Stassi, Diane L.
Summers Jr., Richard G.
Ruan, Xiaoan
Pereda-Lopez, Ana
Kakavas, Stephan J.

(ii) TITLE OF THE INVENTION: NOVEL POLYKETIDE DERIVATIVES
AND RECOMBINANT METHODS FOR MAKING SAME

(iii) NUMBER OF SEQUENCES: 34

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Abbott Laboratories
(B) STREET: 100 Abbott Park Rd.
(C) CITY: Abbott Park
(D) STATE: Illinois
(E) COUNTRY: USA
(F) ZIP: 60064-3500

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE: 16-MAY-1979
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Dianne Casuto
(B) REGISTRATION NUMBER: P-40,943
(C) REFERENCE/DOCKET NUMBER: 4952.US.P2

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (847)-938-3137
(B) TELEFAX: (847)-938-2623
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 925 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGGCCGCTGG	CGGTGATGTT	CACCGGACAG	GGCTCCCAAC	GCCCCGGCAT	GGGACGACAG	60
TTGTACGAGC	ACTTCCCGT	CTTCGCCAG	GCACTGGACG	AGGTCTTCGC	ACTCGCCACC	120
CCCGGACTAC	CGAGGTGAT	TTTCGACCCC	GACCAGCCG	AAACACTCCA	ACGCACCGAC	180
CACGCCAGA	TCGCCCTGTT	CGCCTTCGAA	ACCGCCCTCT	ACCGACTCTG	GGAATCCTGG	240
GGCCTGCGAC	CCGACATGGT	CTGCGACAC	TCGGTGGAG	AAATCACCGC	AGCCCACGTC	300
TCCGGCACCC	TCACCCCTCCC	CGACGCCGTC	CACCTCGTCA	CCACACGGG	CACCCCTCATG	360
CAAACACTGC	CCCCCGGGG	CGCCATGCTC	GGCGTCGCCA	CCGACCCCCA	CACCCCTCCAA	420
CCCCACCTCG	ACAACCACCA	CGACACCATC	TCCATCGCCG	CCATCAACGG	CCCCCACGCC	480
ACCGTCTCT	CCGGCGACCG	CACCACCTC	CACCACATCG	CCACCCAACT	CAACACCAAA	540
CCCTTCACCA	CCACCCCTCAA	CACCCCTCACC	CACCACCCCC	CACACACACC	CCTCATCAGC	600
ATGCTCACCG	CCACACCCAC	CCACCCGAC	ACCACCCACT	GGACCCAGCA	CATCACCGCA	660
CCCCTCCGCT	ACACCGACAC	CCTCCACCAC	CTCCACCACC	ACGGCATTAC	CACCTACCTC	720
GAAATCGGCC	CCGACACCCAC	CCTCACCGCC	CTCGCCCGCA	CCACCCCTCCC	CACCACCCACC	780
CACCTCATCC	CCACCAACCCG	CGCAACAC	AACGAAGTCC	GCAGCACGAA	CGAGGCCTTG	840
GGCAGGGTGT	TCAGCGTGGG	CCACTCGGTG	GACTGGCGGG	CCCTCACTCC	GACCAGGGAGG	900
CGTACCTCCC	TGCGGACGTA	CCCCCT				925

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1030 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CCTAGGACGG	CAGTCCTGCT	CACCGGGCAG	GGTCCCAGC	GTCAGGGCAT	GGGGCGCGAA	60
CTGTACGACC	GGTCACCGGT	GTTCGCCGCC	TCGTTGACG	CGATCTGCGC	TCAACTCGAC	120
GGGCAACTGC	CTCGTCCCCT	CAAGGACGTT	CTCTTCGCCC	CCGAGGGGTC	GGAGGACGCC	180
GCGCTCATCG	ACCGTACGGT	GTTCACACAG	GCGGCTCTGT	TCGCCGTGGA	GACCTCCCTG	240
TTCCGGCTGT	TCGAGGCCA	CGGCCTCGTC	CCGACTTAC	TCATCGGCCA	CTCCATCGGC	300
GAAGTGACCG	CGGCCACCT	GGCCGGGTC	CTCGATCTGG	CGGACGCGTG	CGTCCTGGTC	360
GCCCACCGCG	GCCGCTGTAT	GCAGTCGGCC	CGGGCCGGCG	GGCGATGCG	CGCGGTCCAG	420
GCGAGCGAGG	ACGAGGTACG	CGAGGCCCTC	GCGACCTTCG	ACGATGCGGT	TGCCGTGGCC	480
GGAGTCAACG	GCCCCAACGC	CACCGTCGTC	TCCCGCGACG	AGGACGCGGT	CGAGCGGCTG	540
GTCGCGCGCT	GGCGCGAGCA	GGGCAGGGCG	ACGAAGCGGC	TGCCGGTCAG	CCACGCCCTTC	600
CACTCGCCCGC	ACATGGACGG	GATCGTCGAC	GAGTTCTGTC	CCGCCGTCTC	CGGGCTCACC	660
TTCCGGCTCCC	CGACGATCCC	GGTCGTCCTC	AACGTCACCG	GGACCCCTCGC	CACCGTCGAC	720
CAGCTGACCT	CGCCCCCGTA	CTGGGCACCG	CACATCCGCG	AGGCCGTGCG	CTTCGCCGAC	780
GGGGTGCAGGT	ACCTGGAGGG	CGAGGGCGTC	ACCGAATGGC	TGGAGCTCGG	GCCCGACGGC	840

GTTCCTCGTCG	CCCTGGTCGA	GGACTGCCTG	CGGAAGGAGG	CGGGATCGCT	CGCGTCCGCC	900
CTGCGCAAGG	GGGCGAGCGA	GCCCCACACC	GTGGGCGCGG	CCATGGCCCG	CGCGGTGCTG	960
CGCGGATCCG	GCCCCGACTG	GGCGGCGGTG	TTCCCCGGCG	CACGGGGGT	CGACCTTCCG	1020
ACGTATGCAT						1030

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATCTACACST	CSGGCACSAC	SGGCAAGCCS	AAGGG	35
------------	------------	------------	-------	----

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CTSAAGGCSG	GC GGCGCSTA	CGTSCCSATC	GACCC	35
------------	-------------	------------	-------	----

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGCGAATTCC	TAGGCTGGCG	GTGATGTTCA		30
------------	------------	------------	--	----

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
GCCGGATCCA TGCATACGTC GGCAGGGAGG TAC	33
(2) INFORMATION FOR SEQ ID NO:7:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 28 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single	
(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:	
GCTCGAATTC GCTGGTCGCG GTGCACCT	28
(2) INFORMATION FOR SEQ ID NO:8:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 32 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single	
(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:	
GACGGATCCG GCCCTAGGCT GCGCCCGGCT CG	32
(2) INFORMATION FOR SEQ ID NO:9:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 30 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single	
(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:	
TTGGGATCCT ATGCATTCCA GCGCGAGCGC	30
(2) INFORMATION FOR SEQ ID NO:10:	
(i) SEQUENCE CHARACTERISTICS:	
(A) LENGTH: 26 base pairs	
(B) TYPE: nucleic acid	
(C) STRANDEDNESS: single	
(D) TOPOLOGY: linear	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:	

GAGAAGCTTG GCGCGACTTG CCCGCT

26

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTTTTTAAGC TTGGTACCTG CTCACCGGCA ACACCG

36

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTTTTTGGAT CCCTGCAGCC TAGGGTCGGA GGCACTGCCG GT

42

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTTTTTCTGC AGTATGCATT CCAGGGCAAG CGGTTCT

37

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTTTTTGAAT TCACCGCGTTG CCCGCAGCGT AGGCGC

36

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GATCGAATTC CCTAGGACGG CAGTCCTGCT CACC

34

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GATCGGATCC ATGCATACGT CGGAAGGTCTG ACCCG

35

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTCGAAGAAT TCCCTAGGGT TGCCTTCCTG TTCGAC

36

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TTCGAAAAGC TTATGCATAG ACCGGCAGAT CCACCG

36

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

19

CGGTSAAAGTC SAACATCGG

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

20

GCRATCTCRC CCTGCGARTG

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

44

GAGAGAGGAA CCAACGCGCA CGTGATCGTC GAAGAGGCAC CAGC

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

45

GAGAGAGGAT CCGACCTAGG CGCGGAGGTC ACCGGCGCGA CGGCG

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GAGAGACCTA GGAAGCCGGT GTTCGTGTTG CCCGCCAGG GCT

43

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GAGAGAGGGAT CCGAGGCCGG CCGTGCCTCC GGACCGAAGA CCGCCTC

47

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GAGAGAATTTC CCTAGGGTCG CCTTCGTCTT TCCCGGGCAG G

41

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TTGAGATCTT ATGCATACGA GGGAAAGCCGGC ACCCTGC

37

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 37 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TTGAGATCTT ATGCATACGA GGGAAAGCGGC ACCCTGC

37

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TTGAGATCTT ATGCATACGA GGGAAAGCGGC ACCCTGC

37

(2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1010 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GGCGACCGTG	TCGTGTTCGT	GTTCCCCGGC	CAGGGCTCGC	AGTGGGCCGG	AATGGCCGAG	60
GGGCTGCTGG	AGCGGTCCGG	CGCGTTCCGG	AGTGCGCCGG	ACTCGTGCAG	CGCCGCGCTG	120
CGGCCGTACC	TCGGCTGGTC	GGTGCTGAGC	GTGCTGCGCG	GGGAACCGGA	CGCGCCCTCG	180
CTCGACCGGG	TCGACGTCGT	GCAGCCGGTG	CTGTTACGA	TGATGGTCTC	GCTCGCGCG	240
GTCTGGCGTG	CGCTGGGGGT	GGAACCGGGC	GGCGTGTGCG	GGCACTCGCA	GGGTGAGATC	300
GCCGCTGCC	ATGTCGCCGG	TGCGCTGTGCG	CTGGACGACT	CGGCCCGGAT	CGTCGCCCTG	360
CGCAGTCGGG	CGTGGCTCGG	ACTGGCGGGC	AAGGGCGGC	TGGTGGCGGT	GCGATGCCG	420
GCGGAGGAGC	TGCGGCCGCG	GCTGGTGACG	TGGGGGGACC	GTCTGGCCGT	CGCCGCCCGTC	480
AACAGCCCCG	GTTCTGCGC	CGTCGCAGGC	GACCCGGAGG	CGCTGGCCGA	ACTGGTGGCG	540
CTGCTGACCG	GTGAGGGGGT	GCACGCCCGG	CCGATCCCCG	GGCTCGACAC	GGCGGGCCAC	600
TCGCCGCAGG	TGGACCGT	GGGGCTCAT	CTGCTGGAGG	TGCTGGCCCC	GGTCGCC	660
CGACCGCCG	ACATCCC	GTACTCGACG	GTGACCGCG	GGCTGCTGGA	CGGCACCGAG	720
CTGGACGCCA	CGTACTGGTA	CCGCAACATG	CGCGAGCCCG	TGGAGTTGGA	CGGGGCCACA	780
CGGGCGCTGA	TCGCCGACGG	GCACGACGTC	TTCTGGAGA	CGAGCCCGCA	TCCCATGCTG	840
GCCGTGGCGC	TGGAGCAGAC	GGTCACCGAC	GCGGGCACCG	ACCGGGCGGT	GCTCGGGACC	900
CTGCGCCGCC	GCCACGGCGG	TCCTCGCGCG	CTGGCCCTGG	CCGTCTGCCG	CGCCTTCGCG	960
AGGCGGTCTT	CGGTCCGGGC	GCACGGCCCG	TGGAGTTGCC	CACCTATCCG		1010

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1035 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CGCGCGCCTG	CCTTCGTCTT	TCCCAGGAG	GGCGCCAGT	GGGCCCCACT	GGGAGCGCGG	60
CTCCTCGCGG	ACTCCCCCGT	CTTCGCGCC	AGGGCGAGG	CATGCGCGCG	GGCGCTGGAG	120
CCTCACCTCG	ACTGGTCGGT	CTCTGACGTG	CTGGCCGGCG	CCCCGGGCAC	CCCTCCCATC	180
GACCGGGCCG	ACGTGGTGC	GCGGTGCTG	TTCACCACGA	TGGTCTCGCT	GGCCGCCCTC	240
TGGGAGGCC	ACGGGGTGC	GCGGCCCG	GTCGTGGGCC	ACTCCCAGGG	CGAGGTGGCC	300
GCGGCCTGCG	TGGCCGGTGC	CCTGTCGGT	GACGACGCTG	CCCTGGTGT	CGCCGGACGC	360
AGCAGGCTGT	GGGGGGCGGT	GGCCGGGAAC	GGCAGGATGC	TGCGGGTGT	GGCTCCGGCC	420
GAGCGGATCC	GTGAGCTGCT	CGAACCATGG	CGGCAGCGGA	TTTCGGTGGC	GGCGGTCAAT	480
GGCCCCCGCT	CGGTACCGT	CTCCGGTGAC	GCGCTCGC	TGGAGGAGTT	CGGCGCGCGG	540
CTCTCCGCCG	AGGGGGTGT	GCGCTGGCG	CTGCGGGCG	TGACTTCGC	CGGCCACTCG	600
CCGCAGGTGG	AGGAGTTCCG	CGCTGAGCTC	CTGGACCTGC	TCTCCGGCGT	ACGGCCGGCT	660
CCTTCGCGGA	TACCTTCTT	CTCCACCGTG	ACGGCGGGTC	CTTGCGGCGG	CGACCAGCTG	720
GACGGGGCGT	ACTGGTACCG	CAACACGCC	GAACCCGTGG	AGTTTCGACCC	CACGGTCCGG	780
GCGCTGCTGC	GTGCGGGCCA	TCACACGTT	ATCGAGGTGCG	GTCCGCATCC	GCTGCTCAAC	840
GCCCGCGATCG	ACGAGATCGC	AGCGGACGAG	GGGGTAGCGG	CCACGGCCCT	GCATACGCTC	900
CAGCGGGCG	CTGGCGGCCT	TGACCGCGTG	CGCAACGCGG	TGGGCGCCGC	TTTCGCGCAC	960
GCTGCTCGGG	TCGACTGGAA	CGCCCTGTT	CAGGGCACCG	GTGCGCGCAG	GGTGCCGCTT	1020
CCCTCGTACG	CCTTC					1035

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 328 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Gly	Pro	Leu	Ala	Val	Met	Phe	Thr	Gly	Gln	Gly	Ser	Gln	Arg	Pro	Gly
1					5					10				15	
Met	Gly	Arg	Gln	Leu	Tyr	Glu	His	Phe	Pro	Val	Phe	Ala	Gln	Ala	Leu
					20					25				30	
Asp	Glu	Val	Phe	Ala	Leu	Ala	Thr	Pro	Gly	Leu	Arg	Glu	Val	Met	Phe
							35			40			45		
Asp	Pro	Asp	Gln	Ala	Glu	Thr	Leu	Gln	Arg	Thr	Asp	His	Ala	Gln	Ile
							50			55			60		
Ala	Leu	Phe	Ala	Phe	Glu	Thr	Ala	Ieu	Tyr	Arg	Leu	Trp	Glu	Ser	Trp
							65			70			75		80
Gly	Leu	Arg	Pro	Asp	Met	Val	Cys	Gly	His	Ser	Val	Gly	Glu	Ile	Thr
							85			90			95		
Ala	Ala	His	Val	Ser	Gly	Thr	Leu	Thr	Leu	Pro	Asp	Ala	Val	His	Leu

100	105	110	
Val Thr Thr Arg Gly Thr Leu Met Gln Asn Leu Pro Pro	Gly Gly Ala		
115	120	125	
Met Leu Ala Val Ala Thr Asp Pro His Thr Leu Gln Pro His	Leu Asp		
130	135	140	
Asn His His Asp Thr Ile Ser Ile Ala Ala Ile Asn Gly Pro	His Ala		
145	150	155	160
Thr Val Leu Ser Gly Asp Arg Thr Thr Leu His His Ile Ala	Thr Gln		
165	170	175	
Leu Asn Thr Lys Thr Asn Trp Leu Asn Val Ser His Ala Phe	His Ser		
180	185	190	
Pro Leu Met Gln Pro Ile Leu Gln Pro Phe Thr Thr Thr	Leu Asn Thr		
195	200	205	
Leu Thr His His Pro Pro His Thr Pro Leu Ile Ser Met Leu	Thr Ala		
210	215	220	
Thr Pro Thr His Pro Asp Thr Thr His Trp Thr Gln His Ile	Thr Ala		
225	230	235	240
Pro Val Arg Tyr Thr Asp Thr Leu His His Leu His His Gly	Ile		
245	250	255	
Thr Thr Tyr Leu Glu Ile Gly Pro Asp Thr Thr Leu Thr Ala	Leu Ala		
260	265	270	
Arg Thr Thr Leu Pro Thr Thr His Leu Ile Pro Thr Thr Arg	Arg		
275	280	285	
Asn His Asn Glu Val Arg Ser Thr Asn Glu Ala Leu Gly Arg	Val Phe		
290	295	300	
Ser Val Gly His Ser Val Asp Trp Arg Ala Leu Thr Pro Thr	Gly Arg		
305	310	315	320
Arg Thr Ser Leu Pro Thr Tyr Pro			
325			

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 343 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Pro Arg Thr Ala Val Leu Leu Thr Gly Gln Gly Ser Gln Arg	Gln Gly		
1	5	10	15
Met Gly Arg Glu Leu Tyr Asp Arg Ser Pro Val Phe Ala Ala	Ser Phe		
20	25	30	
Asp Ala Ile Cys Ala Gln Leu Asp Gly Gln Leu Pro Arg Pro	Leu Lys		
35	40	45	
Asp Val Leu Phe Ala Pro Glu Gly Ser Glu Asp Ala Ala Leu	Ile Asp		
50	55	60	
Arg Thr Val Phe Thr Gln Ala Ala Leu Phe Ala Val Glu Thr	Ser Leu		
65	70	75	80
Phe Arg Leu Phe Glu Ala His Gly Leu Val Pro Asp Tyr Leu	Ile Gly		

85	90	95	
His Ser Ile Gly Glu Val Thr Ala Ala His	Leu Ala Gly Val Leu Asp		
100	105	110	
Leu Ala Asp Ala Cys Val Leu Val Ala His	Arg Gly Arg Leu Met Gln		
115	120	125	
Ser Ala Arg Ala Gly Gly Ala Met Ala Ala	Val Gln Ala Ser Glu Asp		
130	135	140	
Glu Val Arg Glu Ala Leu Ala Thr Phe Asp	Asp Ala Val Ala Val Ala		
145	150	155	160
Gly Val Asn Gly Pro Asn Ala Thr Val Val	Ser Gly Asp Glu Asp Ala		
165	170	175	
Val Glu Arg Leu Val Ala Arg Trp Arg Glu	Gln Gly Arg Arg Thr Lys		
180	185	190	
Arg Leu Pro Val Ser His Ala Phe His Ser	Pro His Met Asp Gly Ile		
195	200	205	
Val Asp Glu Phe Val Thr Ala Val Ser Gly	Leu Thr Phe Arg Ser Pro		
210	215	220	
Thr Ile Pro Val Val Ser Asn Val Thr Gly	Thr Leu Ala Thr Val Asp		
225	230	235	240
Gln Leu Thr Ser Pro Ala Tyr Trp Ala Arg	His Ile Arg Glu Ala Val		
245	250	255	
Arg Phe Ala Asp Gly Val Arg Tyr Leu Glu	Gly Val Thr Glu		
260	265	270	
Trp Leu Glu Leu Gly Pro Asp Gly Val Leu	Val Ala Leu Val Glu Asp		
275	280	285	
Cys Leu Ala Lys Glu Ala Gly Ser Leu Ala	Ser Ala Leu Arg Lys Gly		
290	295	300	
Ala Ser Glu Pro His Thr Val Gly Ala Ala	Met Ala Arg Ala Val Leu		
305	310	315	320
Arg Gly Ser Gly Pro Asp Trp Ala Ala Val	Phe Pro Gly Ala Arg Arg		
325	330	335	
Val Asp Leu Pro Thr Tyr Ala			
340			

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 344 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Ala Asp Arg Val Val Phe Val Phe Pro	Gly Gln Gly Ser Gln Trp Ala		
1	5	10	15
Gly Met Ala Glu Gly Leu Leu Glu Arg Ser	Gly Ala Phe Arg Ser Ala		
20	25	30	
Ala Asp Ser Cys Asp Ala Ala Leu Arg Pro	Tyr Leu Gly Trp Ser Val		
35	40	45	
Leu Ser Val Leu Arg Gly Glu Pro Asp Ala	Pro Ser Leu Asp Arg Val		

50	55	60
Asp Val Val Gln Pro Val	Leu Phe Thr Met Met	Val Ser Leu Ala Ala
65	70	75
Val Trp Arg Ala Leu Gly Val Glu Pro Ala Ala	Val Val Gly His Ser	80
85	90	95
Gln Gly Glu Ile Ala Ala Ala His Val Ala Gly Ala	Leu Ser Leu Asp	
100	105	110
Asp Ser Ala Arg Ile Val Ala Leu Arg Ser Arg Ala	Trp Leu Gly Leu	
115	120	125
Ala Gly Lys Gly Gly Met Val Ala Val Pro Met Pro	Ala Glu Glu Leu	
130	135	140
Arg Pro Arg Leu Val Thr Trp Gly Asp Arg Leu Ala	Val Ala Ala Val	
145	150	155
Asn Ser Pro Gly Ser Cys Ala Val Ala Gly Asp Pro	Glu Ala Leu Ala	160
165	170	175
Glu Leu Val Ala Leu Leu Thr Gly Glu Gly Val His	Ala Arg Pro Ile	
180	185	190
Pro Gly Val Asp Thr Ala Gly His Ser Pro Gln Val	Asp Ala Leu Arg	
195	200	205
Ala His Leu Leu Glu Val Leu Ala Pro Val Ala Pro	Arg Pro Ala Asp	
210	215	220
Ile Pro Phe Tyr Ser Thr Val Thr Gly Gly Leu Leu	Asp Gly Thr Glu	
225	230	235
Leu Asp Ala Thr Tyr Trp Tyr Arg Asn Met Arg Glu	Pro Val Glu Phe	240
245	250	255
Glu Arg Ala Thr Arg Ala Leu Ile Ala Asp Gly His	Asp Val Phe Leu	
260	265	270
Glu Thr Ser Pro His Pro Met Leu Ala Val Ala Leu	Glu Gln Thr Val	
275	280	285
Thr Asp Ala Gly Thr Asp Ala Ala Val Leu Gly Thr	Leu Arg Arg Arg	
290	295	300
His Gly Gly Pro Arg Ala Leu Ala Leu Ala Val Cys	Arg Ala Phe Ala	
305	310	315
His Gly Val Glu Val Asp Pro Glu Ala Val Phe Gly	Pro Gly Ala Arg	320
325	330	335
Pro Val Glu Leu Pro Thr Tyr Pro		
340		

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 345 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: None

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Arg Ala Pro Ala Phe Val Phe Pro Gly Gln Gly Ala	Gln Trp Ala Gly	
1	5	10
Leu Gly Ala Arg Leu Leu Ala Asp Ser Pro Val Phe	Arg Ala Arg Ala	15

20	25	30
Glu Ala Cys Ala Arg Ala Leu Glu Pro His Leu Asp Trp Ser Val Leu		
35	40	45
Asp Val Leu Ala Gly Ala Pro Gly Thr Pro Pro Ile Asp Arg Ala Asp		
50	55	60
Val Val Gln Pro Val Leu Phe Thr Thr Met Val Ser Leu Ala Ala Leu		
65	70	75
Trp Glu Ala His Gly Val Arg Pro Ala Ala Val Val Gly His Ser Gln		
85	90	95
Gly Glu Val Ala Ala Ala Cys Val Ala Gly Ala Leu Ser Leu Asp Asp		
100	105	110
Ala Ala Leu Val Ile Ala Gly Arg Ser Arg Leu Trp Gly Arg Leu Ala		
115	120	125
Gly Asn Gly Gly Met Leu Ala Val Met Ala Pro Ala Glu Arg Ile Arg		
130	135	140
Glu Leu Leu Glu Pro Trp Arg Gln Arg Ile Ser Val Ala Ala Val Asn		
145	150	155
Gly Pro Ala Ser Val Thr Val Ser Gly Asp Ala Leu Ala Leu Glu Glu		
165	170	175
Phe Gly Ala Arg Leu Ser Ala Glu Gly Val Leu Arg Trp Pro Leu Pro		
180	185	190
Gly Val Asp Phe Ala Gly His Ser Pro Gln Val Glu Glu Phe Arg Ala		
195	200	205
Glu Leu Leu Asp Leu Leu Ser Gly Val Arg Pro Ala Pro Ser Arg Ile		
210	215	220
Pro Phe Phe Ser Thr Val Thr Ala Gly Pro Cys Gly Gly Asp Gln Leu		
225	230	235
Asp Gly Ala Tyr Trp Tyr Arg Asn Thr Arg Glu Pro Val Glu Phe Asp		
245	250	255
Ala Thr Val Arg Ala Leu Leu Arg Ala Gly His His Thr Phe Ile Glu		
260	265	270
Val Gly Pro His Pro Leu Leu Asn Ala Ala Ile Asp Glu Ile Ala Ala		
275	280	285
Asp Glu Gly Val Ala Ala Thr Ala Leu His Thr Leu Gln Arg Gly Ala		
290	295	300
Gly Gly Leu Asp Arg Val Arg Asn Ala Val Gly Ala Ala Phe Ala His		
305	310	315
Gly Val Arg Val Asp Trp Asn Ala Leu Phe Glu Gly Thr Gly Ala Arg		
325	330	335
Arg Val Pro Leu Pro Ser Tyr Ala Phe		
340	345	